

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed

FEATURES

Location/Qualifiers
1..1986

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 477 a 465 c 451 g 593 t

ORIGIN

Query Match 100.0%; Score 1961; DB 5; Length 1986;
Best Local Similarity 100.0%; Pred. No.: 0;
Matches 1961; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AGGCTGGCATGCTAGTCGAGAGGGCTGGCAGCCTGAAAGGGAGGGGCGCAGCT 60
14 AGGCTGGCATGCTAGTCGAGAGGGCTGGCAGCCTGAAAGGGAGGGGCGCAGCT 73
61 GGGGATACGGCCCGCGCTGGCGACAGCTGGGAGCTGTGGGCGCGGAGGGGCGG 120
74 GGGGATACGGCCCGCGCTGGCGACAGCTGGGAGCTGTGGGCGCGGAGGGGCGG 133
121 GGGGCGGAGCGCGGAGCTGATATCTGGGAGCAAGAGGAGGAGGAGGAGGAGG 180
134 GGGGCGGAGCGCGGAGCTGATATCTGGGAGCAAGAGGAGGAGGAGGAGGAGG 193
181 GGATCTACTGCTTCAAGTGTCTGCTTCAAGCAAGGTTATGCTGCTCTCTGAGCA 240
194 GGATCTACTGCTTCAAGTGTCTGCTTCAAGCAAGGTTATGCTGCTCTCTGAGCA 253
241 ATCTCTGACTGTGATGAAGAATCACTAGTGGGAGCGGAGGAGGAGGAGGAGG 300
254 ATCTCTGACTGTGATGAAGAATCACTAGTGGGAGCGGAGGAGGAGGAGGAGG 313
301 GAAGGCTTTCAGACTTGGGAATATTCGCCAGCATATGCTATTCCTATGCTTACT 360
314 GAAGGCTTTCAGACTTGGGAATATTCGCCAGCATATGCTATTCCTATGCTTACT 373
361 TTGCTCATGCTGGCGCAGCTGATTCATCATCAAGAAATTCACAACTAATAGATT 420
374 TTGCTCATGCTGGCGCAGCTGATTCATCATCAAGAAATTCACAACTAATAGATT 433
421 GTGTTTACTTTTGGCATGTCTTCTGGCTTTTGAGCTGTATTTGAACTTACTTT 480
434 GTGTTTACTTTTGGCATGTCTTCTGGCTTTTGAGCTGTATTTGAACTTACTTT 493
481 TACAGAGCTGTGTCAGAAATTTGGGTTGAGCTGAGTGAATGATGCTAGCTTCT 540
494 TACAGAGCTGTGTCAGAAATTTGGGTTGAGCTGAGTGAATGATGCTAGCTTCT 553
541 GTTCTCAGCACTGGCATGTTTGTCTCATCATCAGCATTCCTTCTAGTAGCTTCT 600
554 GTTCTCAGCACTGGCATGTTTGTCTCATCATCAGCATTCCTTCTAGTAGCTTCT 613
601 TACACTAGTGTATAGCCATGATGATGATATGACAAAGACTTCCATGCTGTGCT 660
614 TACACTAGTGTATAGCCATGATGATGATATGACAAAGACTTCCATGCTGTGCT 673
661 GAGGTACAGCTGGGGCTATCTTAGGCTGGCAATTCAGTGTGCTTTGTTACCAT 720
674 GAGGTACAGCTGGGGCTATCTTAGGCTGGCAATTCAGTGTGCTTTGTTACCAT 733
721 GCTTTGATTGCTGTGTCATGAACACAGGTGAAGATTCTTCAATTTGCTGCTGAT 780
734 GCTTTGATTGCTGTGTCATGAACACAGGTGAAGATTCTTCAATTTGCTGCTGAT 793
781 GCGCTCATACTATTTTGTGCTGTGTGTGTATGATGATGATGATGATGATGAT 840
794 GCGCTCATACTATTTTGTGCTGTGTGTGTATGATGATGATGATGATGATGAT 853
841 GTGATTCACACACACACATTTTGTATATGCTTTTACTCTCTAGAGCTGATCTT 900
854 GTGATTCACACACACACATTTTGTATATGCTTTTACTCTCTAGAGCTGATCTT 913
901 TATGTCAGAACCTGCTATTTCTATTTAATGATTTGATTTCAATGATGCTG 960

914 TATGTCAGAACCTGCTATTTCTATTTAATGATTTTCAATGATGCTG 973
961 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1020
974 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1033
1021 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1080
1034 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1093
1081 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1140
1094 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1153
1141 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1200
1154 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1213
1201 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1260
1214 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1273
1261 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1320
1274 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1333
1321 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1380
1334 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1393
1381 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1440
1394 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1453
1441 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1500
1454 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1513
1501 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1560
1514 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1573
1561 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1620
1574 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1633
1621 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1680
1634 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1693
1681 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1740
1694 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1753
1741 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1800
1754 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1813
1801 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1860
1814 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1873
1861 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1920
1874 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1933
1921 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1961
1934 TTTGCTTTGGCTCTCTAGTCTTACACTGATCTCTTTATGATATACCTGTCAGAGA 1974

RESULT 2
HSH801892

Mon 2.

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 16:29:28 ; Search time 1382.87 Seconds

(without alignments)
11506.297 Million cell updates/sec

Title: PCT-US02-24490-5

Perfect score: 1225

Sequence: 1 cgtacatgtgaagaagatc.....agtcctctgacatgaaaaa 1225

Scoring table:

IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1994485 seqs, 6494577260 residues

Total number of hits satisfying chosen parameters: 3988970

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

GenEmbl_NCHT:*
1: gb_ba:*
2: gb_in:*
3: gb_ov:*
4: gb_ov:*
5: gb_ov:*
6: gb_ov:*
7: gb_ov:*
8: gb_ov:*
9: gb_ov:*
10: gb_ov:*
11: gb_ov:*
12: gb_ov:*
13: gb_ov:*
14: gb_ov:*
15: gb_ov:*
16: gb_ov:*
17: gb_ov:*
18: gb_ov:*
19: gb_ov:*
20: gb_ov:*
21: gb_ov:*
22: gb_ov:*
23: gb_ov:*
24: gb_ov:*
25: gb_ov:*
26: gb_ov:*
27: gb_ov:*
28: gb_ov:*
29: gb_ov:*
30: gb_ov:*
31: gb_ov:*
32: gb_ov:*
33: gb_ov:*
34: gb_ov:*
35: gb_ov:*
36: gb_ov:*
37: gb_ov:*
38: gb_ov:*
39: gb_ov:*
40: gb_ov:*

and is derived by analysis of t
distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	935.4	76.4	1986	5	AX086428
2	935.4	76.4	1986	8	HSM801892
3	932.8	76.1	2028	8	AF395532
4	904.4	73.8	2055	8	BC009255
5	904.4	73.8	2055	8	AK025498
6	697	56.9	2584	9	BC021791
7	407	33.2	420	5	AX071517
8	374	30.5	383	5	AX069718
9	371.4	30.3	474	5	AX197583
10	371	30.3	372	5	AX208199
11	355.8	29.0	608	5	AX198499
12	282.8	23.1	1000	5	AX079399
13	240.2	19.6	181021	8	AF001781
14	237	19.3	42771	8	AC000381
15	119.8	9.8	163447	8	AL354977
16	100.4	8.2	2327	2	AC008208
17	100.4	8.2	170220	2	AC008208
18	100.4	8.2	184887	2	AC008207
19	100.4	8.2	269518	2	AE003751
20	66	5.4	164796	8	AC009058
21	65.8	5.4	112562	8	AC009334
22	65	5.3	155691	8	AL732327
23	64.8	5.3	155691	8	AC016868
24	64.8	5.3	219622	8	AC009630
25	64.2	5.2	87682	8	AC011396
26	64.2	5.2	216021	8	HUAC004787
27	63.8	5.2	175964	8	HS467K16
28	63.6	5.2	65882	8	AL589734
29	62.8	5.1	156818	8	AC104182
30	62.8	5.1	162887	8	AC097635
31	62.8	5.1	167429	8	AC110771
32	62.6	5.1	207503	8	AC087671
33	62.6	5.1	207519	8	AF268884
34	62	5.1	167276	8	AC025262
35	61.8	5.0	164429	8	AC026366
36	61.8	5.0	169552	8	AC015976
37	61.8	5.0	203395	8	AC013566
38	61.4	5.0	2257	5	AX360769
39	61.4	5.0	86155	8	AL589822
40	61.4	5.0	166700	8	AC013759
41	61.4	5.0	176730	8	AC055839
42	61.4	5.0	185568	8	AC105227
43	61.2	5.0	131546	8	HS14209
44	60.8	5.0	69964	8	HS496C20
45	60.8	5.0	140016	8	AC027323

ALIGNMENTS

RESULT 1
AX086428
LOCUS AX086428
DEFINITION Sequence 380 from Patent WO0112659.
ACCESSION AX086428
VERSION AX086428.1 GI:13275964
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Wiemann S.
TITLE Human dna sequences
JOURNAL Patent: WO 0112659-A 380 22-FEB-2001;
German Human Genome Project (DE)

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.

```

FEATURES
    source          1..1986
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
BASE COUNT      477 a 465 c 451 g 593 t
ORIGIN
Query Match      76.4%; Score 935.4; DB 5; Length 1986;
Best Local Similarity 99.4%; Pred. No. 4.8e-246;
Matches 939; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 281 GTGTGGCAGTTCAGATTAGCCACCCGATATGGCTTACTTGGCTCCCAATATAT 340
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1030 GAGATTTATGTTAGAAATTAAGCCACCCGATATGGCTTACTTGGCTCCCAATATAT 1089
QY 341 TTGTTTAAATTTTCTTCATCCAGCCCTCACAAGAGAGAGATTCTTCTCCCTGTGA 400
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1090 TTGGTTTAAATTTTCTTCATCCAGCCCTCACAAGAGAGAGATTCTTCTCCCTGTGA 1149
QY 401 TCCACTTATATGTCCTGTGGCGCTGTGGCTCTCTGTGCACTTCAGAAATGTTACCACT 460
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1150 TCCACTTATATGTCCTGTGGCGCTGTGGCTCTCTGTGCACTTCAGAAATGTTACCACT 1209
QY 461 TGTGTTAAGATATGCGCTGGAGCACTATCTGTGATGATGATGATGATGATGATGATGAT 520
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1210 TGTGTTAAGATATGCGCTGGAGCACTATCTGTGATGATGATGATGATGATGATGATGAT 1269
QY 521 AACTGCTCTCTGTTGGGCTCTGTGCACTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 580
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1270 AACTGCTCTCTGTTGGGCTCTGTGCACTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1329
QY 581 TCAGGGGCGCTTGTATTTGATCCAGATTTTACGAATTTGCTACAGACCACCATCCA 640
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1330 TCAGGGGCGCTTGTATTTGATCCAGATTTTACGAATTTGCTACAGACCACCATCCA 1389
QY 641 CACGTGCCAGAGAGCACTGTGAATGCTGTGGGAAAGAGTGATGATGATGATGATGATGATGAT 700
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1390 CACGTGCCAGAGAGCACTGTGAATGCTGTGGGAAAGAGTGATGATGATGATGATGATGATGAT 1449
QY 701 CAGCAGCTCTCTTCTGTGCAATTTGCACTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 760
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1450 CAGCAGCTCTCTTCTGTGCAATTTGCACTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1509
QY 761 TCAGTTACCAAAACCTTTTGCAGAGAGAGCTGTGGCCACCCGGATGTTCTTCTACTGACAT 820
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1510 TCAGTTACCAAAACCTTTTGCAGAGAGAGCTGTGGCCACCCGGATGTTCTTCTACTGACAT 1569
QY 821 GAATGACAGAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 880
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1570 GAATGACAGAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1629
QY 881 AGTGGATTGGACACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 940
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1630 AGTGGATTGGACACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1689
QY 941 AGATGATCAGCTTGGCTTATAGACATTTCTTGATGCTTCTAGATCTTCAAAAGTGTCT 1000
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1690 AGATGATCAGCTTGGCTTATAGACATTTCTTGATGCTTCTAGATCTTCAAAAGTGTCT 1749
QY 1001 GCGGGGATTCATGCTCCCTCTGTCAGATCACTATATACATGATGATGATGATGATGATGAT 1060
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1750 GCGGGGATTCATGCTCCCTCTGTCAGATCACTATATACATGATGATGATGATGATGATGAT 1809
QY 1061 CCTCAAAACCCCGAAGCAAAAGCAATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1120
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1810 CCTCAAAACCCCGAAGCAAAAGCAATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1869
QY 1121 TGGCCCCAAGAGACACATTTGTTAACTATGATTCAGTACGATGATGATGATGATGATGATGAT 1180
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1870 TGGCCCCAAGAGACACATTTGTTAACTATGATTCAGTACGATGATGATGATGATGATGATGAT 1929
QY 1181 TCATGCGCTGATCACTTGAATTAAGGCTTCTGACATGAGAAAA 1225
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Db 1930 TCATGCGCTGATCACTTGAATTAAGGCTTCTGACATGAGAAAA 1974
RESULT 2
LOCUS      HSM801892
DEFINITION Homo sapiens mRNA; cDNA DKFZp586m2420 (from clone DKFZp586m2420);
complete cds.
ACCESSION AL36927
VERSION   AL36927.1 GI:12053348
KEYWORDS
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 1986)
            Mambuti, R., Heubner, D., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
            Direct Submission
            Submitted (12-MAR-2002) MIPS, Am Klopferspitz 18a, D-82152
            Martinsried, GERMANY
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ). Email: s.wiemann@dkfz-heidelberg.de;
            sequenced by AGO (Berlin/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            This clone (DKFZp586m2420) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/proj/cDNA/.
FEATURES
    source          1..1986
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /map="11"
                    /clone="DKFZp586m2420"
                    /tissue_type="uterus"
                    /clone_id="586 (synonym: hute1). Vector pSport1; host
                    DH10B; sites NotI + SalI/MluI"
                    /dev_stage="adult"
                    1..1986
                    /gene="DKFZp586m2420"
                    23..1858
                    /gene="DKFZp586m2420"
                    /note="strong similarity to S.cerevisiae A19p"
                    /codon_start=-1
                    /product="hypothetical protein"
                    /protein_id="CA866861.1"
                    /db_xref="GI:12053349"
                    /translation="MASRGARORLKGSGASSGDTAPADKRLRELLSREAGAEHFR
                    LSGNKAGOVAPGEGSTAFKCLLSARICALLLSINISQDEFENYMEPTHYLYEGG
                    VCKRRFLHSRMLATLVSTGFCSSALPSSFCMYTTLIMTGMWMDKTSIAVIG
                    LVAGAILGWFSALGALPLAFLDLVYKHKRSEFMSLNAALFLVAVVVIDSYVYK
                    LVAPLNLVLYNVTFGHDPDLYGTEPMYFLINGFLNFVAFLALVPLTSLMEYL
                    LQRFVQNLGHPRWMLTAPMYIWFILFQPHREERFLFPVYPLICGVALSALOK
                    CYHFVFORLREHYTYSNMLAGVLEFLGLSFSVALFRGHGRLDLYPEFYIA
                    TDPTIHTVEGRPVNVYCVKEMVREYRSLDLPDMQIOLPSEFROQLKPPAEGLA
                    TRVPTDMDQNLSEPSRIDISKRYLVDLDMTRTPREPKTSSNKEEWISLAIKPF
                    LDASRSSKLRLAIVPFLSDQYTVVYVNTILKPKAKQIRKSSG"
                    1950..1955
                    /gene="DKFZp586m2420"
                    1966
                    /gene="DKFZp586m2420"
    polyA_site
    polyA_signal
    BASE COUNT      477 a 465 c 451 g 593 t
ORIGIN
Query Match      76.4%; Score 935.4; DB 8; Length 1986;
Best Local Similarity 99.4%; Pred. No. 4.8e-246;
Matches 939; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 16:29:28 : Search time 993.409 Seconds

(without alignments)
11506.297 Million cell updates/sec

Title: PCT-US02-24490-3

Perfect score: 880

Sequence: 1 cgtacaatgtgaagaagtc.....cagtaaatgcatatttag 880

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1994485 segs, 6494577260 residues

Total number of hits satisfying chosen parameters: 3988970

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBml_NOHTG:*
1: gb_ha:*
2: gb_in:*
3: gb_ov:*
4: gb_ov:*
5: gb_ov:*
6: gb_ov:*
7: gb_ov:*
8: gb_ov:*
9: gb_ov:*
10: gb_ov:*
11: gb_ov:*
12: gb_ov:*
13: gb_ov:*
14: gb_ov:*
15: gb_ov:*
16: gb_ov:*
17: gb_ov:*
18: gb_ov:*
19: gb_ov:*
20: gb_ov:*
21: gb_ov:*
22: gb_ov:*
23: gb_ov:*
24: gb_ov:*
25: gb_ov:*
26: gb_ov:*
27: gb_ov:*
28: gb_ov:*
29: gb_ov:*
30: gb_ov:*
31: gb_ov:*
32: gb_ov:*
33: gb_ov:*
34: gb_ov:*
35: gb_ov:*
36: gb_ov:*
37: gb_ov:*
38: gb_ov:*
39: gb_ov:*
40: gb_ov:*

and is derived by analysis of th

SD

tribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	580.4	66.0	1985	5	AX086428
2	580.4	66.0	1985	8	HSMB01892
3	580.4	66.0	2028	8	AF395532
4	549.4	62.4	2055	8	BC009255
5	549.4	62.4	2094	8	AK025498
6	484.4	55.0	2584	9	BC021791
7	341	38.8	383	5	AX069718
8	326	37.0	420	5	AX071517
9	282.8	32.1	1000	5	AX079399
10	164.8	18.7	181021	8	AP001781
11	130.6	14.8	163447	8	AL354977
12	74.4	8.5	474	5	AX197583
13	74	8.4	372	5	AX208199
14	73.4	8.3	608	5	AX198499
15	65.8	7.5	112562	8	AC099334
16	65.6	7.5	164796	8	AC009058
17	65	7.4	190325	8	AL732327
18	64.8	7.4	155691	8	AC016868
19	64.8	7.4	219622	8	AC009630
20	64.2	7.3	87692	8	AC011396
21	63.8	7.3	175964	8	HS467K16
22	63.8	7.2	216021	8	AL589734
23	63.2	7.2	65882	8	AL589734
24	62.4	7.1	156818	8	AC104182
25	62.4	7.1	162887	8	AC097635
26	62.2	7.1	167429	8	AC110771
27	62	7.0	207503	8	AC087671
28	62	7.0	207519	8	AF286884
29	61.4	7.0	2257	5	AX360769
30	61.4	7.0	166700	8	AC013759
31	61.4	7.0	185568	8	AC105227
32	61.2	7.0	162276	8	AC025262
33	61.2	7.0	169552	8	AC015976
34	61	6.9	163042	8	AC016245
35	61	6.9	176730	8	AC055839
36	60.8	6.9	69964	8	HS496C20
37	60.8	6.9	86155	8	AL589822
38	60.8	6.9	164429	8	AC026366
39	60.8	6.9	203395	8	AC073596
40	60.4	6.9	179350	8	AC009337
41	60.2	6.8	87637	8	AL390960
42	60.2	6.8	140016	8	AC027323
43	60	6.8	37661	5	AX147034
44	60	6.8	162208	8	AC073321
45	60	6.8	171998	8	AC026120

ALIGNMENTS

RESULT 1
AX086428
LOCUS AX086428
DEFINITION Sequence 380 from Patent WO011659.
ACCESSION AX086428
VERSION
KEYWORDS
SOURCE
ORGANISM human.
Human sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 186)
Wiemann, S.
Human dna sequences
Patent: WO 011659-A 380 22-FEB-2001;
German Human Genome Project (DE)

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,

FEATURES
 source
 location/Qualifiers
 1. 1986
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 477 a 465 c 451 g 593 t
 ORIGIN

Query Match 66.0%; Score 580.4; DB 5; Length 1986;
 Best Local Similarity 98.8%; Pred. No. 1.4e-144;
 Matches 595; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

280 GTGTGTCAGATTCAGATTTAGGCGCAGCTATGCTTACCTGGCTCCAAATGATAT 339
 1030 GAGATTTTCAGTTTCAGATTTAGGCGCAGCTATGCTTACCTGGCTCCAAATGATAT 1089
 340 TTGGTTTAAATTTTCTTCATCCAGCTTCACAAAGAGAGATTTCTTTCCCTGTGA 399
 1090 TTGGTTTAAATTTTCTTCATCCAGCTTCACAAAGAGAGATTTCTTTCCCTGTGA 1149
 400 TTCACTTATATGCTCTGTGCGCTGTGCTCTCTGCACTTCAGAAATGTTACCACT 459
 1150 TTCACTTATATGCTCTGTGCGCTGTGCTCTCTGCACTTCAGAAATGTTACCACT 1209
 460 TGTGTTCAACGATATCGCTGAGAGCACTATCTGATGACATCGAATGCTGCGATTAG 519
 1210 TGTGTTCAACGATATCGCTGAGAGCACTATCTGATGACATCGAATGCTGCGATTAG 1269
 520 AACTGCTTCTCTGTTGGGCTCTTTCATTTTCTGCTCTGTGCGACTGTTACAGAGATA 579
 1270 AACTGCTTCTCTGTTGGGCTCTTTCATTTTCTGCTCTGTGCGACTGTTACAGAGATA 1329
 580 TCACGGGCCCCCTGATTTGATCCAGAAATTTTACGAATGCTTCAGACCAACCATCA 639
 1330 TCACGGGCCCCCTGATTTGATCCAGAAATTTTACGAATGCTTCAGACCAACCATCA 1389
 640 CACTGTCAGAGAGCAGACCTGTGAATGCTGTGTGGGAAAAGAGTGTATGATTTCC 699
 1390 CACTGTCAGAGAGCAGACCTGTGAATGCTGTGTGGGAAAAGAGTGTATGATTTCC 1449
 700 CACAGCTTCTCTCTCCATCAATTTGAGAGCTTCATTCATTCAGACAGAGG 759
 1450 CACAGCTTCTCTCTCCATCAATTTGAGAGCTTCATTCATTCAGACAGAGG 1509
 760 TCAGTTACCAAAACCTTTTCAGAGAGCCTCTGCGCACCAGATTTCTCTACATGACAT 819
 1510 TCAGTTACCAAAACCTTTTCAGAGAGCCTCTGCGCACCAGATTTCTCTACATGACAT 1569
 820 GAATGACGAGATTC-AGAAGAGCCATCCAGATATATGATATCAGTAATGCCATTATTT 878
 1570 GAATGACGAGATTC-AGAAGAGCCATCCAGATATATGATATCAGTAATGCCATTATTT 1629
 879 AG 880
 1630 AG 1631

RESULT 2
 HSM801892 1986 bp mRNA linear PRI 20-MAR-2002
 LOCUS Homo sapiens mRNA; cDNA DKFZp586M2420 (from clone DKFZp586M2420);
 DEFINITION complete cds.
 ACCESSION AL136927
 VERSION AL136927.1 GI:12053348
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 1986)
 AUTHORS Mambelli, R., Heubner, D., Meves, H. W., Gassenhuber, J. and Wiemann, S.
 TITLE Direct Submission
 JURNAL Submitted (12-MAR-2002) MIPS, Am Klopferplatz 18a, D-82152
 Martinsried, GERMANY

COMMENT

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ). Email: s.wiemann@dkfz-heidelberg.de; sequenced by AGORA (Berlin/Germany) within the CDNA sequencing consortium of the German Genome Project.
 This clone (DKFZp586M2420) is available at the R2PD in Berlin. Please contact the R2PD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@r2pd.de. Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cdna/>.

FEATURES

source

location/Qualifiers
 1. 1986
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="11"
 /clone="DKFZp586M2420"
 /issue_type="uterus"
 /clone_lib="586 (synonym: hute1). Vector pSport1; host DH10B; sites NotI + SalI/MluI"
 /dev_stage="adult"
 1. 1986
 /gene="DKFZp586M2420"
 23. 1936
 /note="strong similarity to S.cerevisiae Al99p"
 /product="hypothetical protein"
 /protein_id="CAB66861.1"
 /db_xref="GI:12053349"
 /db_xref="SEPRMBL:094068"
 /translation="NASSRGARORLKGSGASGSDTAPADRLKRLGSRBAGAEHRT
 LSGNRAGQVWAEQSTAFKCLSLARLALSNISDCDETFYWEPTHYLIAGEFOT
 WEKSPAYAIRSYAYLLHAWPAHFHRILOTNKLIVFYLRLCLLAVSCICEFYKA
 VCKKGLHYSRMALFVLYSTGMCSSSFSSRCMTYLLAGMWDKTSIVAG
 VAAALILGHPFSAALGLPAPDLVLYKHKRPFHMSLALLFLVRYVYDYSYK
 LVIAPLNLYIVNFTFPHGDLTGTEFETITLNGELNENVAFLALVLPISMETL
 LQRFVQNLGHPYWLTLAPVYIMFIIIFLQPKKEKEFLPVPLCLGVALSALQK
 CYHFEQRLRLEHYTVTSNMAIGLTVLEGLSFSKSAFLFGYHGPDLPEFYRIA
 TPTIITVPEGRPVNVCKEWMYREFSSFLPDNMOLQIPESEFROLKPPAEGLA
 TRIVPDKNDONLLEPSRYIDISKCHYLVDLTMRTPREPRYSKKEEMISLAYRPF
 LDASSSKLLRAFYVPELSDQTVVYNYTILMPKAKOIRKKSQSG"

gene

CDS

polyA_signal
 /gene="DKFZp586M2420"
 1966
 /gene="DKFZp586M2420"
 polyA_site
 /gene="DKFZp586M2420"
 BASE COUNT 477 a 465 c 451 g 593 t
 ORIGIN

Query Match 66.0%; Score 580.4; DB 8; Length 1986;
 Best Local Similarity 98.8%; Pred. No. 1.4e-144;
 Matches 595; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

280 GTGTGTCAGATTCAGATTTAGGCGCAGCTATGCTTACCTGGCTCCAAATGATAT 339
 1030 GAGATTTTCAGTTTCAGATTTAGGCGCAGCTATGCTTACCTGGCTCCAAATGATAT 1089
 340 TTGGTTTAAATTTTCTTCATCCAGCTTCACAAAGAGAGATTTCTTTCCCTGTGA 399
 1090 TTGGTTTAAATTTTCTTCATCCAGCTTCACAAAGAGAGATTTCTTTCCCTGTGA 1149
 400 TTCACTTATATGCTCTGTGCGCTGTGCTCTCTGCACTTCAGAAATGTTACCACT 459
 1150 TTCACTTATATGCTCTGTGCGCTGTGCTCTCTGCACTTCAGAAATGTTACCACT 1209
 460 TGTGTTCAACGATATCGCTGAGAGCACTATCTGATGACATCGAATGCTGCGATTAG 519
 1210 TGTGTTCAACGATATCGCTGAGAGCACTATCTGATGACATCGAATGCTGCGATTAG 1269
 520 AACTGCTTCTCTGTTGGGCTCTTTCATTTTCTGCTCTGTGCGACTGTTACAGAGATA 579
 1270 AACTGCTTCTCTGTTGGGCTCTTTCATTTTCTGCTCTGTGCGACTGTTACAGAGATA 1329
 580 TCACGGGCCCCCTGATTTGATCCAGAAATTTTACGAATGCTTCAGACCAACCATCA 639
 1330 TCACGGGCCCCCTGATTTGATCCAGAAATTTTACGAATGCTTCAGACCAACCATCA 639